

# ARIMA 模型在海南省 2014—2017 年居民自杀死亡率中的应用

刘莹 胡锡敏 陈言 符振旺

570203 海口,海南省疾病预防控制中心热带病与慢性病预防控制所地方病与慢性病预防控制室

通信作者:陈言, Email:jkban702@163.com

DOI: 10.3760/cma.j.issn.0254-6450.2018.05.024

**【摘要】** 目的 分析海南省居民自杀报告死亡率变化趋势,探索时间序列模型在自杀死亡率预测中的应用。**方法** 应用时间序列模型拟合海南省 2014 年 1 月至 2016 年 12 月自杀死亡率资料,并预测该地区 2017 年 1—6 月的自杀死亡率。**结果** 海南省 2014 年 1 月至 2017 年 6 月该地区报告自杀死亡 576 人,平均死亡率为 1.5/10 万;所建立的 ARIMA 模型较好地拟合了既往时间序列的自杀死亡率,预测值的动态趋势与实际值基本一致,2017 年 1—6 月的实际值均在预测值 95% 的可信区间范围内。**结论** 确定海南省自杀死亡率的时间序列模型为“ARIMA(0, 1, 0)(0, 0, 0)<sub>12</sub>”,且模型的预测效果较好,可以用来预测海南省因自杀导致的死亡情况。

**【关键词】** 自杀; 伤害; 时间序列

## Application of ARIMA model in prediction of mortality rate of suicide in Hainan province

Liu Ying, Hu Ximin, Chen Yan, Fu Zhenwang

Office of Endemic and Chronic Disease Control, Department of Tropical and Non-communicable Disease Control and Prevention, Hainan Provincial Center for Disease Control and Prevention, Haikou 570203, China  
Corresponding author: Chen Yan, Email: jkban702@163.com

**【Abstract】** **Objective** To analyze the trend of suicide mortality in residents of Hainan province, and explore the application of time sequence model in the prediction of the mortality of suicide. **Methods** The mortality data of suicide in residents of Hainan province between January, 2014 and December, 2016 were collected and analyzed with time sequence model and the mortality rate of suicide during January-June, 2017 in Hainan was predicted with the model. **Results** During January, 2014 to June 2017, a total of 576 suicide cases were reported in Hainan, the mortality rate was 1.5/100 000. The established ARIMA model had good fitting for the suicide mortality in previous times and the prediction result was quite similar to the actual mortality, the predicted mortality rate was within the 95% confidence interval of the actual rate. **Conclusion** The time sequence model for the prediction of suicide mortality in Hainan was “ARIMA (0, 1, 0) (0, 0, 0)<sub>12</sub>”, and the prediction effect of the model was better, which can be used to predict the suicide mortality in Hainan.

**【Key words】** Suicide; Injury; Time sequence

自杀行为对家庭、社区和整个国家,并对生者产生影响。2017 年 9 月 10 日世界预防自杀日,WHO 发布的信息显示,每年有近 80 万人死于自杀,还有更多的人试图自杀,并且 78% 的全球自杀发生在低收入和中等收入国家,在普通人群中,自杀企图是自杀的唯一危险因素。更重要的是,自杀发生在人的整个生命周期,是 2015 年全球 15~29 岁人群死亡的第二大原因<sup>[1]</sup>。十几年前,已有报道国内每年约 28.7 万人死于自杀<sup>[2]</sup>,海南地区尚无全人群的自杀数据分析报告,仅有大学生自杀意念等的调查<sup>[3]</sup>,调查结果显示海南医学生的自杀意念发生率为

15.4%。故了解海南省全人群的自杀死亡趋势对相关机构制定现在和未来的自杀预防政策尤为重要,本研究利用 ARIMA 模型对海南省的自杀死亡率进行建模和预测,提示其未来趋势。

### 资料与方法

1. 资料来源:采用海南省 2014—2017 年的自杀监测数据,来源于海南省 CDC 死因监测系统。所有原始数据来源于由卫生部门统一印发的《居民死亡医学证明(推断)书》,在医疗机构发生的死亡个案由医疗机构负责填写;在非医疗机构死亡的个案由社

区卫生服务中心(乡镇/街道卫生院)、村卫生室(社区卫生服务站)医生通过面访家属和知情人后填写。两者均在24 h内向本单位疾控科上报,由疾控科医生进行审核,确保所有数据真实可靠。死因分类采用国际疾病分类第10版(ICD-10),本研究选取的对象为故意自害,即根本死因在X60~X84之间的所有个案。常住人口数来自海南省统计局。根据自杀个案和人口数计算自杀死亡率。

2. 时间序列分析:首先将所有数据按照年月格式转换为时间序列,并分为两部分:2014年1月至2016年12月和2017年1-6月,前半部分作为建模数据,后半部分作为验证数据。利用时间序列图、自相关函数图和偏自相关函数图来判断原始数据是否随机、平稳和季节性;拟定p、d、q值和P、D、Q值后通过差分、移动平均或季节性平稳转换等方法将序列数据变换处理成平稳序列后再用时间序列自回归移动平均结合模型(auto-regressive integrated moving average, ARIMA)来分析处理,手动调整参数选取最优的ARIMA(p, d, q)(P, D, Q)模型;最后生成预测图,用后半部分数据的拟合值是否在实际值的可信区间范围来检验模型的匹配度。

3. 统计学分析:采用SPSS 22.0软件对数据进行模型的建立、分析和预测。

### 结果

1. 自杀致死概况:2014年1月至2017年6月海南省共报告576例因自杀导致死亡案例,各年度死亡率之间的差异有统计学意义( $\chi^2=60.009, P<0.0001$ ),2016年最高,2017年最低(表1)。从图1可以看出,自杀死亡率的时间序列不平稳,呈现波动趋势,季节性波动不明显,2016年的自杀死亡率明显高于其他年份。除去2016年,2014-2017年的自杀死亡率整体呈下降趋势。整体来看,最高值位于

表1 海南省2014-2017年自杀死亡人数

年份	人口数	自杀死亡人数	死亡率(10万)
2014	8 952 659	132	1.47
2015	8 952 405	98	1.10
2016	9 024 281	223	2.47
2017 <sup>a</sup>	11 364 455	123	1.08
合计	38 293 799	576	1.50

注:<sup>a</sup>2017年数据为1-6月

2016年5月(3.56/10万),最低值位于2015年5月(0.53/10万)。

### 2. 时间序列模型:

(1)识别模型:绘制原始序列的序列图(图1)、自相关函数(auto-correlation function, ACF)图及偏自相关函数(partial auto-correlation function, PACF)图(图2)以识别原序列的随机性、平稳性和季节性,可见原始序列在短时滞上,自相关系数大且为正,而偏自相关系数在时滞为1时有极大值,时滞数>1后偏自相关系数骤减,提示原始序列为非平稳序列,需要进行平稳化处理。经自然对数转化、差分后序列变得相对平稳(图3)。

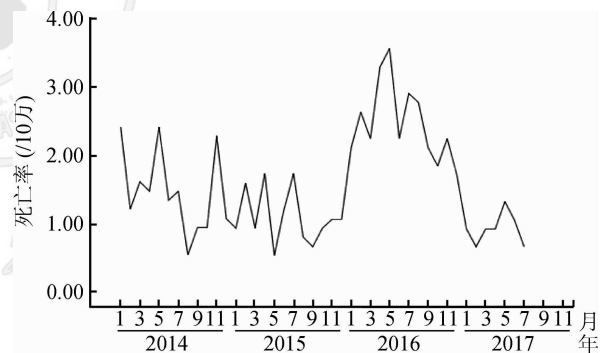


图1 海南省2014-2017年自杀死亡率时间序列图

(2)对模型参数进行估计:模型的参数采取从低阶到高阶的顺序进行尝试,阶数一般不超过2,因此

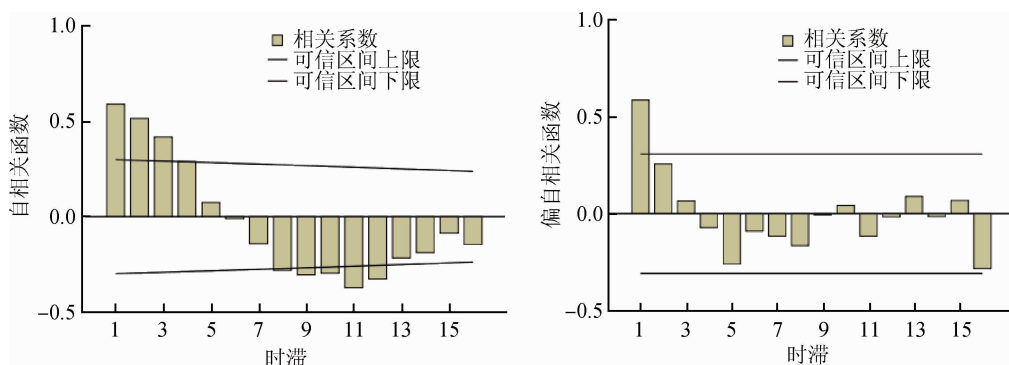


图2 原始序列的自相关函数及偏自相关函数

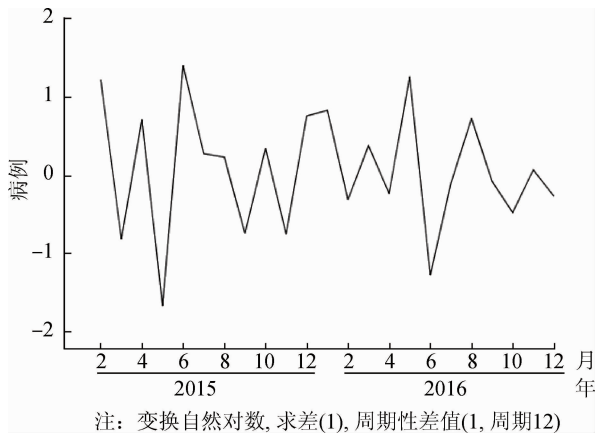


图3 经自然对数转化、差分后的序列图

在 2 阶以下逐个检验其 BIC、 $R^2$  等值, 其中 BIC 值越小,  $R^2$  越大, 表明模型预测效果越好<sup>[4]</sup>。根据 BIC、 $R^2$  等参数值选取较优的 4 种模型进行对比(表 2); 之后, 以 2017 年 1—6 月的实际值与 4 种模型预测值的差距作为观测值进行 Ljung-Box Q 检验, 确保残差(考察模型假设的合理性及数据的可靠性)彼此独立, 故在 4 种模型中选了 Sig 值最小的一个, 即  $(0, 1, 0)(0, 0, 0)_{12}$  为预测模型。

表 2 4 种模型的参数值

模型	$R^2$	BIC	Sig
$(0, 1, 0)(0, 0, 0)_{12}$	0.119	-0.509	0.696
$(1, 0, 0)(0, 0, 0)_{12}$	0.312	-0.633	0.813
$(0, 1, 1)(0, 0, 0)_{12}$	0.296	-0.602	0.989
$(0, 1, 2)(0, 0, 0)_{12}$	0.301	-0.477	0.984

(3) 预测: 图 4 为所取模型预测 2017 年海南省自杀死亡率的效果图, 从图中可以看出预测值和实际值的吻合较好, 所有实际值均落在预测值的 95% 可信区间内, 说明该模型预测效果好。

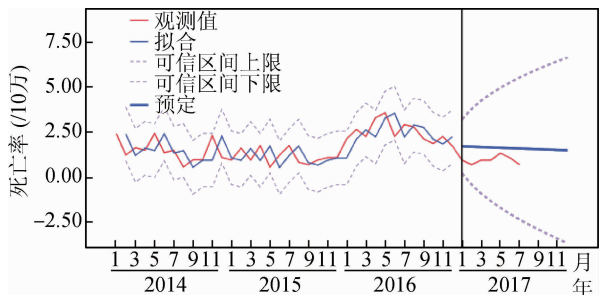


图4 模型  $(0, 1, 0)(0, 0, 0)_{12}$  预测 2017 年海南省自杀死亡率效果图

### 讨 论

自杀引起国家疾病负担和对人群期望寿命的影

响很大。我国因自杀造成的 YLLs/YLDs (过早死亡损失寿命年/伤残损失寿命年) 为 90.03<sup>[5]</sup>。自杀的最大风险因素是自杀未遂, 目前为止, 只有少数国家将自杀预防纳入健康重点, 仅 28 个国家报告有国家自杀预防战略。要将自杀预防纳入当地的健康重点, 首先了解当地的自杀风险, 但在全球范围内, 自杀和自杀企图的数据可用性和质量都很差。故如何利用有限的对自杀死亡率进行科学预测, 具有十分重要的意义。

目前在公共卫生领域, 应用较多的统计预测方法主要有回归分析和时间序列分析两大类。一般的线性回归模型只是简单的线性关系变化模型, 对随机性、波动性较大的数据拟合效果较差<sup>[6]</sup>, 预测精度欠理想。并且, 回归分析通常都需要考虑各种对预测对象的影响因素, 分析这些因素对预测对象造成的影响, 而众多的影响因素和某些数据资料的不易得到导致回归分析在实际应用中非常受限。时间序列分析克服了回归分析的限制, 利用事物发展的延续性建立模型, 以时间综合替代了趋势、季节、人为等影响因素, 但并不要求明确所有的影响因素。ARIMA 模型预测法适用于各种类型时间序列数据的预测, 可对时间序列进行具体识别, 初步选定一个试用模型后再用一系列统计方法来检验该模型的适用性, 在这个过程中也可以对模型进行必要的调整。ARIMA 通过 ACF 及 PACF 图等方法可以判定序列是否平稳<sup>[7]</sup>, 适用于存在周期性波动的疾病研究<sup>[8]</sup>。王伟辉等<sup>[9]</sup>认为 ARIMA 建模法可将时间序列的特征分析及预处理、模型识别与定阶、模型参数估计与检验、残差检验等过程整合, 大大降低了模型应用的难度。国内很多学者已经将此方法成功应用于流感<sup>[10]</sup>、疟疾<sup>[11]</sup>、手足口病<sup>[12]</sup>、心血管病<sup>[13]</sup>、甲型肝炎<sup>[14]</sup>等疾病预测方面。并且, 使用 SPSS 软件的方法也比较成熟<sup>[15]</sup>。本研究以时间代替趋势、季节等影响因素, 系统构建了 ARIMA 时间序列预测模型, 并且预测值均落在实际观测值的可信区间内, 说明模型预测结果理想, 由于该模型是动态模型, 今后继续收集补充自杀死亡信息将进一步提高模型的稳健度, 增加预测精度, 延长预测周期。

据 2014 年的 WHO 报道, 2012 年有 804 000 人死于自杀, 其中约 75% 发生在发展中国家和低收入国家, 而中国自杀人数约占世界范围内自杀人数的 22%<sup>[16]</sup>。从图 1 可以看出, 海南省 2014—2017 年的自杀趋势在 2015 年 9 月以后呈明显的上升趋势, 但于 2016 年 11 月后又重新回落, 是因为 2015 年海南省进

行了一次较大规模的漏报调查,调查后一段时期内死亡监测工作有一个较为明显的改进,提示相关部门应加强质量控制,避免漏报。总体来说,海南省2014—2017年的自杀死亡率明显低于2010年全国自杀死亡率(10.15/10万)<sup>[17]</sup>,也低于辽宁省2003年的自杀死亡率(4.7/10万)<sup>[18]</sup>,但我国卫生部统计结果显示,自杀已经成为15~34岁人群死亡的首要因素<sup>[19]</sup>,而2016年北京市职业高中生的自杀报告率仅为2.5%<sup>[20]</sup>。据此推断,全人群中自杀致死的应该更少。国内有研究表明,自杀好发于秋冬季,以冬季尤为明显<sup>[21]</sup>,海南省处于热带地区,全年温度较高,自杀死亡率低可能也与此有关。有研究表明,南亚各国的自杀死亡率在0.43/10万~331.0/10万之间不等<sup>[22]</sup>,并且自杀死亡率有逐年降低的趋势,国内也有较多研究发现该趋势<sup>[23-24]</sup>。从预测图来看,除2016年有个小幅度的上升外,总体趋势在下降,尤其是预测的2017年,处于稳步下降的趋势,但自杀死亡率下降并不代表自杀意念或自杀行为的减少,仍应不断加强自杀防御体系的建设。

本文的局限性在于海南省常规监测工作起步较晚,数据虽然容易获得,但存在数据质量的问题,如数据不完整等;另外,由于观测值较少(<50),序列中波动的产生原因有待进一步研究。

综上所述,ARIMA模型可作为医疗卫生机构合理安排自杀的干预工作的警示性预报。目前我国自杀问题研究主要集中在流行病学方面,应当开展多学科、跨学科的研究,为预防和处理自杀提供更多的科学依据。

利益冲突 无

### 参 考 文 献

- [1] World Health Organization. Preventing suicide: a resource for media professionals[R]. Geneva: WHO, 2017.
- [2] Phillips MR, Li XY, Zhang YP. Suicide rates in China, 1995–1999 [J]. *Lancet*, 2002, 359 (9309): 835–840. DOI: 10.1016/S0140-6736(02)07954-0.
- [3] 吴玲,王小丹,田治.海南医学生自杀意念的现况调查[J]. *中国健康心理学杂志*, 2013, 21 (10): 1587–1588. DOI: 10.13342/j.cnki.cjhp.2013.10.027.  
Wu L, Wang XD, Tian Y. The cross-sectional survey of suicidal intention in medical students in Hainan [J]. *Chin J Health Psychol*, 2013, 21 (10): 1587–1588. DOI: 10.13342/j.cnki.cjhp.2013.10.027.
- [4] 李娜,殷菲,李晓松.时间序列分析在结核病发病预测应用中的初步探讨[J]. *现代预防医学*, 2010, 37(8): 1426–1428.  
Li N, Yin F, Li XS. Study on the time series analysis in forecasting the incidence of tuberculosis in Sichuan province [J]. *Mod Prev Med*, 2010, 37(8): 1426–1428.
- [5] 高欣,王临虹,金叶,等.1990年与2013年中国人群自杀疾病负担分析[J]. *中华流行病学杂志*, 2017, 38 (10): 1325–1329. DOI: 10.3760/cma.j.issn.0254-6450.2017.10.006.  
Gao X, Wang LH, Jin Y, et al. Disease burden caused by suicide in the Chinese population, in 1990 and 2013 [J]. *Chin J Epidemiol*, 2017, 38 (10): 1325–1329. DOI: 10.3760/cma.j.issn.0254-6450.2017.10.006.
- [6] 张欣.南昌市气象因素与冠心病急性发作关系的研究[D].南昌:南昌大学,2013:5.  
Zhang X. Study on the relationship between meteorological factors and coronary heart disease acute attack of Nanchang [D]. Nanchang: Nanchang University, 2013: 5.
- [7] 明亮,徐德忠,林明和,等.应用时间序列模型预测疟区疟疾发病率[J]. *第四军医大学学报*, 2004, 25 (6): 507–510. DOI: 10.3321/j.issn:1000-2790.2004.06.008.  
Wen L, Xu DZ, Lin MH, et al. Prediction of malaria incidence in malaria epidemic area with time series models [J]. *J Fourth Mil Med Univ*, 2004, 25 (6): 507–510. DOI: 10.3321/j.issn:1000-2790.2004.06.008.
- [8] 刘雷,张连生,汤恒,等.ARIMA乘积季节模型在丙肝发病预测中的应用[J]. *中华疾病控制杂志*, 2014, 18(4): 366–367.  
Liu L, Zhang LS, Tang H, et al. Application of ARIMA multiplicative seasonal model on Hepatitis C incidence [J]. *Chin J Dis Control Prev*, 2014, 18(4): 366–367.
- [9] 王伟辉,崔燕,韩宝泉.某医院普外科住院病例时间序列分析[J]. *中国卫生统计*, 2017, 34(2): 296–297.  
Wang WH, Cui Y, Han BQ. Time series analysis on hospitalized cases of department of general surgery in a hospital [J]. *Chin J Health Stat*, 2017, 34(2): 296–297.
- [10] 戴亚欣,蒲柳艳,李科峰,等.舟山市2009—2015年流感样病例发病趋势时间序列分析[J]. *中国预防医学杂志*, 2017, 18(2): 147–149. DOI: 10.16506/j.1009-6639.2017.02.016.  
Dai YX, Pu LY, Li KF, et al. Time series analysis on the trend of influenza-like cases in Zhoushan in 2009–2015 [J]. *Chin Prev Med*, 2017, 18 (2): 147–149. DOI: 10.16506/j.1009-6639.2017.02.016.
- [11] 李白坤,王建军,武松,等.基于遥感技术和时间序列分析法预测皖北疟疾的发病率[J]. *中华疾病控制杂志*, 2017, 21 (3): 291–294. DOI: 10.16462/j.cnki.zhjbkz.2017.03.018.  
Li BK, Wang JJ, Wu S, et al. Study on the prediction of malaria incidence in the northern Anhui Province based on remote sensing techniques and time series analysis [J]. *Chin J Dis Control Prev*, 2017, 21 (3): 291–294. DOI: 10.16462/j.cnki.zhjbkz.2017.03.018.
- [12] 王伶,姚文清.利用时间序列模型分析预测辽宁手足口病疫情趋势[J]. *中国卫生统计*, 2016, 33(5): 847–849.  
Wang L, Yao WQ. A time series model was used to analyze and predict the epidemic situation of hand foot and mouth disease outbreaks in liaoning province [J]. *Chin J Health Stat*, 2016, 33 (5): 847–849.
- [13] 钱旭君,沈月平,贺天锋,等.宁波市大气颗粒物与人群因心脑

- 血管疾病死亡的时间序列研究[J]. 中华流行病学杂志, 2016, 37(6): 841-845. DOI: 10.3760/cma.j.issn.0254-6450.2016.06.020.
- Qian XJ, Shen YP, He TF, et al. Relationship between particulate matters and cardio-cerebrovascular mortality in Ningbo: a time-series study[J]. Chin J Epidemiol, 2016, 37(6): 841-845. DOI: 10.3760/cma.j.issn.0254-6450.2016.06.020.
- [14] 陶长余, 沉郁, 章士军. 时间序列分析在甲型肝炎疫情预测中的应用[J]. 医学动物防制, 2016, 32(10): 1113-1115. DOI: 10.7629/yxdwfwz201610017.
- Tao CY, Chen Y, Zhang SJ. Application of time series analysis in forecasting the epidemic trend of hepatitis A [J]. J Med Pest Control, 2016, 32(10): 1113-1115. DOI: 10.7629/yxdwfwz201610017.
- [15] 王瑜, 刘光文, 贾蕾. 和田市2010—2014年肠道传染病流行特征时间序列分析[J]. 中国公共卫生, 2016, 32(9): 1265-1267. DOI: 10.11847/zgggws2016-32-09-35.
- Wang Y, Liu GW, Jia L. Epidemiological characteristics of intestinal infectious diseases in Hetian city, 2010-2014: a time series analysis [J]. Chin J Public Health, 2016, 32(9): 1265-1267. DOI: 10.11847/zgggws2016-32-09-35.
- [16] Sun L, Zhang J. Potential years of life lost due to suicide in China, 2006-2010 [J]. Public Health, 2015, 129(5): 555-560. DOI: 10.1016/j.puhe.2015.02.012.
- [17] 蔡钥, 胡楠, 刘韞宁, 等. 中国人群2010年自杀死亡现状分析[J]. 中国预防医学杂志, 2012, 13(6): 480-482. DOI: 10.16506/j.1009-6639.2012.06.009.
- Cai Y, Hu N, Liu WN, et al. Analysis on the current situation of suicide deaths in China in 2010 [J]. Chin Prev Med, 2012, 13(6): 480-482. DOI: 10.16506/j.1009-6639.2012.06.009.
- [18] 杨晓丽, 那军, 于连政, 等. 1992—2003年辽宁省城市居民自杀死亡率趋势分析[J]. 中国慢性病预防与控制, 2008, 16(2): 200-201. DOI: 10.3969/j.issn.1004-6194.2008.02.033.
- Yang XL, Na J, Yu LZ, et al. Analysis on urban residents suicide mortality trends in Liaoning province in 1992-2003 [J]. Chin J Prev Contr Chron Non-Commun Dis, 2008, 16(2): 200-201. DOI: 10.3969/j.issn.1004-6194.2008.02.033.
- [19] 赵军燕, 邢淑芬. 国外青少年的自杀风险因素研究及启示[J]. 首都师范大学学报: 社会科学版, 2014(1): 151-156. DOI: 10.3969/j.issn.1004-9142.2014.01.022.
- Zhao JY, Xing SF. A study of risk factors for foreign adolescent suicide and its implications [J]. J Cap Norm Univ: Soc Sci Ed, 2014(1): 151-156. DOI: 10.3969/j.issn.1004-9142.2014.01.022.
- [20] 张艺阳, 段佳丽, 吕若然, 等. 北京市职业高中生自杀行为特征及其影响因素分析[J]. 中国学校卫生, 2016, 37(2): 224-228, 232. DOI: 10.16835/j.cnki.1000-9817.2016.02.020.
- Zhang YY, Duan JL, Lyu RR, et al. Characteristics of suicide behaviors among occupational high school students and associated factors [J]. Chin J Sch Health, 2016, 37(2): 224-228, 232. DOI: 10.16835/j.cnki.1000-9817.2016.02.020.
- [21] 高慧, 王冬, 丛林. 院前急救中自杀患者的临床特点及性别差异[J]. 临床急诊杂志, 2016, 17(12): 952-956. DOI: 10.13201/j.issn.1009-5918.2016.12.016.
- Gao H, Wang D, Cong L. The clinical features and gender differences of emergency patients with self-inflicted injuries [J]. J Clin Emerg (China), 2016, 17(12): 952-956. DOI: 10.13201/j.issn.1009-5918.2016.12.016.
- [22] Jordans MJD, Kaufman A, Brenman NF, et al. Suicide in South Asia: a scoping review [J]. BMC Psychiatry, 2014, 14: 358. DOI: 10.1186/s12888-014-0358-9.
- [23] 王妍敏, 杨晓明, 沈冰, 等. 上海市静安区1975—2014年户籍人口自杀死亡情况分析[J]. 中国公共卫生, 2016, 32(3): 369-372. DOI: 10.11847/zgggws2016-32-03-32.
- Wang YM, Yang XM, Shen B, et al. Suicide mortality among permanent residents in Jing'an district of Shanghai, 1975-2014 [J]. Chin J Public Health, 2016, 32(3): 369-372. DOI: 10.11847/zgggws2016-32-03-32.
- [24] 陈亦晨, 李小攀, 杨琛, 等. 2002—2013年上海市浦东新区居民自杀死亡流行特征及趋势分析[J]. 中国健康教育, 2017, 33(2): 174-177. DOI: 10.16168/j.cnki.issn.1002-9982.2017.02.021.
- Chen YC, Li XP, Yang C, et al. Analysis on the epidemiological characteristics and trend of the suicide data of residents in Pudong New Area of Shanghai, 2002-2013 [J]. Chin J Health Educ, 2017, 33(2): 174-177. DOI: 10.16168/j.cnki.issn.1002-9982.2017.02.021.

(收稿日期: 2017-10-31)

(本文编辑: 王岚)